How Might Financial Market Information Be Used for Supervisory Purposes?*

John Krainer

Economist
Federal Reserve Bank of San Francisco

Jose A. Lopez

Senior Economist
Federal Reserve Bank of San Francisco

Bank supervisory monitoring, both on-site and off-site, generates a wealth of information with which to judge the safety and soundness of banks and bank holding companies (BHCs). For BHCs with publicly traded securities, the monitoring efforts of investors generate additional information that may complement the supervisory information set. In this paper, we address three public policy questions related to how supervisors might use this financial market information. First, can financial markets detect changes in BHC risk characteristics? To address this question, we summarize the academic literature on the topic and present our own empirical results using BHC stock returns and bond spreads. We find that securities prices signal changes in supervisory ratings of BHC condition up to a year prior to their assignment. Second, do securities prices provide information that complements supervisory information? Using forecasts generated by an off-site monitoring model developed by Krainer and Lopez (2001), we find that securities prices do improve forecasts of supervisory ratings changes, although the improvement is not statistically significant. Third, what is an appropriate level of accuracy to demand of financial market signals and off-site monitoring models more generally? We examine this question by studying the model's ratio of correct forecasts to incorrect forecasts.

1. Introduction

Banks are thought to play a special role in the economy as monitors of investment projects and aggregators of borrower information. Banks also provide households and firms with sources of liquidity and allegedly play a role in the transmission of monetary policy. Without exception, countries with modern economies have developed banking sectors to help finance real activity. But the types of functions that banks carry out—taking deposits from households and investing them in illiquid and difficult-to-evaluate projects—necessarily expose them to risks of failure. For example, economic shocks that raise depositor demand for additional liquidity can cause banks' funding to dry up and potentially lead to decreased bank lending or the inefficient liquidation of existing projects.

To guard against such systemic risks, countries have created public safety nets and established banking supervisory authorities of various forms. In the United States, deposit insurance and the Federal Reserve's discount window are intended to protect both depositors and banks from systemic shocks. In addition, multiple government agencies supervise and regulate the banking sector to ensure that it is not overly vulnerable to systemic shocks.

Currently, most bank supervisory agencies conduct on-site examinations to monitor the health of banking institutions and look closely at bank financial conditions and operational processes. However, it is increasingly obvious that a banking institution's condition can deteriorate rapidly, and, if banks are examined relatively infrequently, then supervisory assessments can become outdated quickly.[1] Thus, supervisors also have developed off-site

*We thank Rob Bliss for sharing his bank holding company debt database with us. We thank Fred Furlong and Simon Kwan for extensive comments and suggestions, Anita Todd for editorial assistance, and Judy Peng, Paul Schwabe, Ryan Stever, and Chishen Wei for research assistance.

1. Note that for large, complex banking organizations, the Federal Reserve now maintains a continuous supervisory presence. See DeFerrari and Palmer (2001) for further discussion.

monitoring methods to augment their on-site exams. These methods are appealing because they provide predictive information that can identify problems at an early stage when it should be easier and cheaper for supervisors to deal with them.

An interesting development in bank supervision has been the attempt to harness financial market data, such as bank holding company (BHC) debt and equity prices, for the monitoring effort. For example, the Basel Committee on Banking Supervision (2001) proposed a new international accord on commercial bank capital requirements, one pillar of which focuses on market discipline of banks. A joint study by the Board of Governors of the Federal Reserve System and the U.S. Department of the Treasury (2000) concluded that subordinated debt issuance could be a way to encourage market discipline at depository institutions. The report further states that the U.S. bank supervisory agencies will continue to use and should enhance their application of financial market data to evaluate current and future BHC conditions.[2]

Financial market participants and the information they produce can aid supervisors in two distinct ways. The debt markets can impose *direct* discipline on banking organizations by demanding high prices for the financing of risky activities. Thus, the appropriate pricing of risk may discourage banks from taking on risk profiles that supervisors would discourage. *Indirect* discipline—potentially more useful to supervisors—refers to the signals derived from the banking organization's traded securities. Investors, like regulators, have much to gain from learning about the condition of a banking institution. Thus, market prices for bank debt and equity should reflect a consensus opinion among investors about the value of the firm. If market prices are accurate, then these prices presumably could give supervisors an alternative, reliable assessment of the condition of a banking institution.

In the U.S., supervisors monitor an array of financial market variables, such as changes in stock prices, expected default probabilities, and spreads on subordinated notes and debentures, as part of their off-site monitoring.[3] Researchers, meanwhile, have been trying to quantify the usefulness of financial market data for the supervisory process and to discover the best ways to use the information. In this paper, we examine this research as it relates to BHCs, which in the United States are supervised by the Federal Reserve.

We do so by addressing three specific policy questions. First, we ask the basic question of whether changes in BHC risk characteristics get incorporated into BHC securities prices. The consensus in the academic literature is that the market pricing of debt and equity securities reflects BHC risk. We extend this literature by presenting our own empirical results that changes in BHC stock returns and bond spreads anticipate (i.e., lead) changes in supervisory BHC ratings.

Second, we ask whether financial market prices can provide complementary information not already in the supervisory information set. We explore this question using an off-site monitoring model proposed by Krainer and Lopez (2001). Using data from the 1990s, we find that incorporating BHC financial market variables into the model improves its in-sample fit. However, including financial market variables in the model does not produce ratings forecasts that are statistically different from forecasts based on supervisory data alone.

While financial markets correctly signal changes in condition that our benchmark model fails to identify, the problem is that markets also incorrectly signal changes in condition that do not occur. This finding leads naturally to our third policy question: What degree of accuracy should supervisors demand of financial market signals and of off-site monitoring models more generally? We examine this question by studying our model's ability to forecast supervisory rating downgrades correctly and weighing the number of correct signals against the number of signals that incorrectly forecast a ratings change. We show that the model that includes financial market variables produces a different set of correct forecasts than the core model without these variables. These differences, we argue, can be exploited since this extended model is relatively better than the core model at forecasting ratings changes for publicly traded BHCs.

In conclusion, we provide empirical evidence on the question of how useful financial market information could be for supervisory off-site monitoring. The evidence suggests that using BHC financial market data certainly should not hinder the monitoring effort. The financial market variables we study do not provide unambiguously better forecasts of BHC condition when combined with supervisory variables in our model. Upon closer inspection, though, we see that financial market variables help identify ratings changes at publicly traded BHCs that otherwise would not be flagged by our benchmark core model. Therefore, it should be beneficial to combine forecasts from the two models, especially since this benefit can be achieved at a fairly low cost.

The paper is organized as follows. In Section 2, we survey the academic literature on the topic. In Section 3, we

2. See also Board of Governors (1999).

3. For a more complete discussion of how Federal Reserve supervisors monitor financial market information, see Supervisory Letters 95-03 and 02-01 (Board of Governors 1995, 2002).

provide a brief summary of the supervision of bank holding companies in the United States. In Section 4, we present our event studies regarding supervisory assessments, known as BOPEC ratings, as they relate to BHC stock returns and debt yields. In Section 5, we examine two sets of BOPEC forecasts generated from different specifications of the model: a core model based on a set of supervisory variables, and an extended model that incorporates both equity and debt market variables into the core model. We also examine the BOPEC forecast accuracy trade-offs implied by these two models. We present our conclusions in Section 6.

2. Literature Review

Many researchers have examined how financial markets anticipate or respond to different types of events. For the purpose of this paper, we are interested in two related questions in the academic literature. First, can financial markets accurately assess the conditions of financial firms, especially BHCs, or are their assets too opaque for investors? Second, in what ways might financial market information be of use for supervisory monitoring?

2.1. Literature on Bank Opacity

Modern theories of financial intermediation, such as Diamond (1984), stress the special role of banks as monitors and processors of information. That is, banks specialize in gathering and monitoring balance sheet information from their borrowers, whose assets otherwise would be too opaque for other institutions to evaluate. An important strand of empirical work that has flowed from this model of banking seeks to determine whether banks, as specialists in solving problems of asymmetric information between firms and investors, are themselves opaque to investors. In other words, can financial market investors accurately assess the value of bank assets given their opacity?

To our knowledge, the first academic paper to explore the relationship between BHC conditions and the responses of their financial market securities was written by Pettway (1976). The author found that BHC debt yields and stock returns were not sensitive to variations in BHC capital ratios in the early 1970s. Subsequent to this work, many studies have examined whether investors in bank securities respond to changes in bank default risk.

With respect to bank debt, the evidence regarding bank opaqueness has been mixed depending on both the type of debt instrument and the sample period, but the overall conclusion is that investors in debt securities do price changes in bank risk. As surveyed by Gilbert, et al. (2001), studies of the market for large, uninsured certificates of deposit (CDs) in the 1970s generally found no evidence that investors in these instruments responded to changes in bank risk. However, studies based on data from the 1980s found that investors responded to such changes. For example, Hannan and Hanweck (1988) found that CD rates were sensitive to the volatility of banks' returns on assets and capital ratios during the 1980s.

Interestingly, the empirical relationship between the yields on bank subordinated notes and debentures, which are typically less senior and longer term than CDs, and measures of bank risk has been harder to establish; see the survey by the Board of Governors (1999). Studies of these debt yields from the mid-1980s by Avery, et al. (1988) as well as Gorton and Santomero (1990) found no evidence of sensitivity to measures of bank risk. However, as argued by Flannery and Sorescu (1996), this supposed failure by investors to price different bank risks could be due to investor perception of a real or implicit government guarantee of bank liabilities. In fact, they found that, following the passage of the 1991 Federal Deposit Insurance Corporation Improvement Act and its stated retreat from a "too-big-to-fail" policy, spreads on traded bank debt securities became more responsive to risk measures. Several other studies, such as Jagtiani, et al. (2000) and Morgan and Stiroh (2001), found similar results for the latter half of the 1990s.[4]

With regard to investors in bank equities, the evidence regarding bank opacity, as derived from many studies, suggests that stock market investors do price changes in bank risk. As summarized by Flannery (1998), this empirical literature provides broad support for the hypothesis that bank equity investors incorporate risk-related information into BHC stock prices. For example, Kho, et al. (2000) found that the degree to which BHC stock prices were affected by currency crises in emerging market economies in the late 1990s depended on the degree of the BHC's exposure to these economies. Furthermore, Dahiya, et al. (2003) found that stock prices of lead lending BHCs fall significantly when a major corporate borrower announced its default or bankruptcy. When the BHCs were ranked according to their exposures to such distressed borrowers, the stock price decline for the low exposure banks was insignificant, while that for the high exposure banks was large and significant.

4. A recent study by Covitz, et al. (2002) found that the bank subordinated debt premiums (over similar maturity Treasury securities) used in these studies are biased due to the risk-sensitive nature of the managerial decision to issue such debt. That is, we can assume that the issuers of bank subordinated debt before and after 1991 were sufficiently different so as to be subject to important sample selection issues. After adjusting for this sample selection bias, the authors found that bank subordinated debt prices were sensitive to changes in bank risk prior to the Federal Deposit Insurance Corporation Improvement Act.

Another branch of the literature regarding bank opacity has sought to determine whether banks are more opaque to investors relative to comparable nonbank firms. Flannery, et al. (2002) used stock market data to address this question, based on the premise that compared to more transparent firms, opaque firms' stocks should exhibit different market microstructure properties. Opaque firms' earnings also should be more difficult to forecast. They found that larger BHCs traded on the New York and American Stock Exchanges were as easy for investors to evaluate as similarly sized nonbanks. In addition, smaller BHCs traded on the NASDAQ were relatively easier to assess than similarly sized nonbanks. They concluded that asset opacity is not a prominent feature of bank holding companies. In contrast, Morgan (2003) found that credit rating agencies were more likely to disagree on debt ratings for financial intermediaries, such as banks and BHCs, and that these disagreements mainly arose from assets, such as loans, that are harder for outside monitors to assess. These seemingly contradictory results could be due to differing incentives facing stock investors and rating agencies or different research methodologies. Further research into this topic is necessary.

2.2. Literature Linking Market and Supervisory Assessments

Given that financial markets have been shown capable of providing information on bank conditions, we now turn to the question of whether this information might be of use to supervisors. The academic literature most related to this paper examines the correlation between market evaluations of bank conditions and supervisory assessments, which are summarized by supervisory ratings. Note that market participants are not actively trying to guess what supervisory ratings will be, nor do they officially learn the outcomes of supervisory inspections. Instead, this literature tests whether there is an alignment between investor and supervisory assessments that could allow supervisors to use market data for their own purposes.

Most studies in this area have examined equity market or debt market information separately, although a few recent studies have examined both. With respect to equity market information alone, Berger and Davies (1998) used an event study framework to examine whether daily BHC stock prices react to changes in their subsidiary banks' supervisory ratings. Even though these ratings are confidential, they found that BHC stock prices respond to these changes, implying that supervisory assessments provide valuable information that the equity market can detect. The authors conjectured that market participants view supervisory ratings as both certifications of bank financial statements and indicators of future regulatory treatment.[5] Several recent studies have provided complementary evidence. Gunther, et al. (2001) and Hall, et al. (2001) found that equity market signals provide useful information that supplements supervisory assessments. Elmer and Fissel (2001) as well as Curry, et al. (2001) further supported this conclusion by finding that equity market variables add value to supervisory models of bank failure.

Similar studies have been conducted using debt market information. DeYoung, et al. (2001) found that supervisory information significantly affects contemporaneous and subsequent changes in the spreads of bank debentures over the corresponding Treasury rates. Specifically, they found that the private information component of bank supervisory ratings affects debt spreads several months after their assignments. Gilbert, et al. (2001) found that default risk premia for jumbo CDs, as derived from financial statement variables, do not predict bank supervisory downgrades as well as a standard early warning model.

Since the interests of bank subordinated debt holders and bank supervisors generally are considered to be aligned, several studies have advocated that subordinated debt prices be incorporated into the supervisory process.[6] Evanoff and Wall (2001) examined this proposition directly by testing the degree to which subordinated debt spreads provide supervisors with additional information. In their study, they modeled changes in the supervisory ratings of banks and BHCs that have outstanding subordinated debt over the period 1990 to 1999 as a function of lagged subordinated debt spreads and regulatory capital ratios. They found that subordinated debt spreads do as well as or better than any of the capital ratios at explaining supervisory rating changes.

A few recent studies have examined the relationship between securities market information, both equity and debt, and bank ratings. Gropp and Richards (2001) found that changes in agency ratings for European banks have little impact on bond prices and a strong effect on equity prices in the case of rating downgrades. Extending this work, Gropp, et al. (2002) examined the ability of equity and debt market variables for European banks to signal changes in

---

5. Note that Allen, et al. (2001) do not find evidence that supervisory ratings affect equity market assessments when the supervisory ratings of bank management actually become public knowledge. However, since these results are tied to the special case of changes in bank charter status, it is not clear how general these results are.

6. Hancock and Kwast (2001) provide support and additional guidance for the use of subordinated debt spreads in supervisory monitoring and highlight difficulties with currently available data sources. Bliss (2000) objects to this viewpoint, arguing that supervisory interests may diverge from bondholder interests in that the parties may disagree on the relative riskiness of different bank portfolios.

bank financial conditions. Using ordered logit models at several horizons and a proportional hazard model, they found that both equity-based measures of bank default and subordinated debt spreads are useful for detecting changes in bank agency ratings. Interestingly, they found that equity-based default measures are less useful when banks are closer to default and that subordinated debt spreads have signal value only close to defaults.

Bongini, et al. (2002) found different patterns among equity and debt market indicators with respect to signaling financial distress at individual Asian banks during the East Asia crisis of 1997. The authors found that information based on stock prices or on agency ratings did not improve upon the historical information contained in balance sheet data. However, stock prices responded more quickly to changing financial conditions than did agency ratings.

Turning to U.S. supervisory BHC ratings, Berger, et al. (2000) found that financial market and supervisory assessments appear to focus on different aspects of BHC performance. Supervisory BHC ratings are most closely related to agency ratings, as indicated by the finding that supervisory ratings Granger-cause agency bond ratings, and vice versa. However, this type of interrelation is not apparent between supervisory ratings and equity market assessments; that is, the authors found very little Granger-causality from equity market variables to supervisory ratings. Furthermore, they found that, after accounting for equity and debt market assessments, supervisory ratings do not contribute substantially to the modeling of future indicators of BHC performance, such as changes in nonperforming loans. Our study is most closely related to this study, although our sample periods and empirical methods differ. Overall, however, both of our studies find that monitoring by supervisors, bond market investors, and equity market investors all produce complementary information on BHC performance.

In summary, the broad consensus from these empirical studies is that financial market investors do conduct reasonable monitoring of BHCs and that the information they generate is different from that generated by supervisors. As shown in Krainer and Lopez (2001, 2003) and as will be shown below, financial market information can complement supervisory monitoring, especially with regard to off-site monitoring models.

3. Supervision of Bank Holding Companies

In the United States, the Federal Reserve is the supervisor of BHCs as well as financial holding companies, which were created by the Gramm-Leach-Bliley Act of 1999. Full-scope, on-site inspections are a key element of the supervisory process. These inspections generally are conducted on an annual basis, particularly in the case of large and complex BHCs.[7]

Although the inspection process also includes limited and targeted inspections that may or may not be conducted on-site, we focus our analysis on full-scope, on-site inspections since they provide the most comprehensive supervisory BHC assessments.

At the conclusion of an inspection, the supervisory team assigns the institution a numerical rating, called a composite BOPEC rating, that summarizes its opinion of the BHC's overall health and financial condition.[8] The BOPEC acronym stands for the five key areas of supervisory concern: the condition of the BHC's **B**ank subsidiaries, **O**ther nonbank subsidiaries, **P**arent company, **E**arnings, and **C**apital adequacy. BOPEC ratings are assigned according to an absolute scale ranging from 1, indicating strong performance, to 5, indicating very poor performance. BOPEC ratings of 1 or 2 indicate that the BHC is not considered to be of supervisory concern. Note that BOPEC ratings are highly confidential and are not made available to the public.

Between on-site inspections, when private supervisory information cannot be gathered as readily, supervisors monitor BHCs using a well-specified off-site monitoring system; see Board of Governors (1995, 2002) for further discussion. Of particular importance to our analysis are the three primary sources of information used in the surveillance process. One source, known as the BHC Performance Report, is a detailed summary of BHC quarterly regulatory reporting forms; see Board of Governors (2001) for a complete description. As of March 1999, the report summarized approximately 800 BHC variables across several years. From this report, certain variables are selected as key performance criteria, and if a BHC fails to meet these criteria in a given quarter, this is noted as a supervisory exception that requires further monitoring.

A second source of information for off-site BHC monitoring is the supervisory CAMELS ratings assigned to banks within a bank holding company. As with BOPEC ratings, CAMELS ratings are assigned after bank examinations and are not made public. The CAMELS acronym refers to six key areas of concern: the bank's **C**apital adequacy, **A**sset quality, **M**anagement, **E**arnings, **L**iquidity, and **S**ensitivity to risk. The composite CAMELS ratings

7. A complex BHC has nonbank subsidiaries that extend credit or have debt outstanding to the general public. See DeFerrari and Palmer (2001) for an overview of the supervisory process for large, complex banking organizations. Note that the frequency of such inspections for BHCs with less than $1 billion in assets has been reduced, as described in Board of Governors (2002).

8. For an overview of supervisory rating systems in the U.S. and the rest of the world, see Sahajwala and Van der Bergh (2000).

also range from 1 to 5, with banks assigned lower ratings for better performance. Since BHC conditions are closely related to the conditions of their subsidiary banks, the off-site BHC surveillance process includes monitoring recently assigned CAMELS ratings.

As with on-site BHC inspections, on-site bank examinations occur at approximately an annual frequency, which is long enough for the gathered supervisory information to decay and become less representative of the bank's condition.[9] To address this issue, the Federal Reserve instituted an off-site monitoring system for banks, known as the System for Estimating Examiner Ratings (SEER), in 1993.[10] The SEER system consists of two separate models, one that forecasts bank failures over a two-year horizon and one that forecasts CAMELS ratings for the next quarter. The model that we are most interested in here is the latter, which is an ordered logit model with five categories corresponding to the five possible values of the CAMELS rating. The model is estimated every quarter in order to reflect the relationship between selected financial ratios and the two most recent quarters of CAMELS ratings. Significant changes in a bank's CAMELS rating as forecasted by the SEER model could be sufficient to warrant closer monitoring of the bank. The off-site BHC surveillance program also explicitly monitors the SEER model's forecasted CAMELS ratings.

A third information source is BHC financial market information, when available. Supervisors monitor BHC securities prices at various horizons. For example, as stated in Board of Governors (1995), supervisory staff at the Federal Reserve Banks are expected to monitor stock prices for BHCs. If a BHC exhibits irregular stock price movements, this can be noted as an exception that requires closer monitoring during the regular surveillance process.

To examine the contribution of BHC financial market information to the supervisory process statistically, we construct a data set that combines BOPEC ratings, BHC variables collected during the supervisory process, and BHC financial market variables.[11] The complete data set spans the period from 1990 to 1999 and consists of 3,399 BOPEC assignments. Note that we will make use of different subsamples of this larger data set, depending on the application, throughout the paper.

To underscore the point that BOPEC ratings are absolute ratings and not relative ratings, Table 1 reveals a clear trend in assigned ratings that is tied to the broader U.S. economy. In 1990, over 30 percent of the BOPEC assignments were ratings of 3 or worse, and just 18 percent of the ratings were 1. However, by 1999, when the economy and the banking sector were in better condition, less than 7 percent of the BOPEC assignments were 3 or worse, and 43 percent of the ratings were 1.

4. Univariate Event Studies

This section addresses the first policy question, namely whether changes in BHC risk characteristics are incorporated into BHC security prices. We examine this question by conducting two sets of event studies to determine whether BHC financial market variables anticipate BOPEC assignments. The event studies focus on firm-specific variations in financial market variables that could be useful to supervisors.

There are several issues that arise in the process of conducting these event studies. First, as mentioned, BOPEC ratings are confidential and are not made public. Hence,

Table 1
BOPEC Ratings in the Sample, 1990–1999

	BOPEC ratings				
	1	2	3	4–5	Total observations
1990	46	135	54	27	262
% of total	17.6	51.5	20.6	10.3	
1991	48	140	76	36	300
% of total	16.0	46.7	25.3	12.0	
1992	55	194	75	52	376
% of total	14.6	51.6	19.9	13.8	
1993	96	216	56	28	396
% of total	24.2	54.5	14.1	7.1	
1994	136	211	32	22	401
% of total	33.9	52.6	8.0	5.5	
1995	143	210	31	18	402
% of total	35.6	52.2	7.7	4.5	
1996	194	195	21	3	413
% of total	47.0	47.2	5.1	0.7	
1997	176	178	16	1	371
% of total	47.4	48.0	4.3	0.3	
1998	113	108	14	1	236
% of total	47.9	45.8	5.9	0.4	
1999	104	122	12	4	242
% of total	43.0	50.4	5.0	1.7	
Total	1,111	1,709	387	192	3,399
%	32.7	50.3	11.4	5.6	

9. See Cole and Gunther (1998) as well as Hirtle and Lopez (1999) for further discussion of this issue.

10. For a complete description of the SEER system, see Cole, et al. (1995). The statistical analysis of supervisory bank ratings within the Federal Reserve System dates back to Stuhr and Van Wicklen (1974); see also Korobow, et al. (1977).

11. Throughout our analysis, we differentiate between "supervisory" variables, which are generated during BHC inspections or from mandatory supervisory reporting, and "financial market" variables derived from securities prices. This terminology unfortunately obscures the fact that supervisors also monitor financial market variables.

equity and debt market variables cannot directly anticipate BOPEC ratings. Instead, we must assume that changes in BHC conditions and investor expectations of their future profitability will lead to changes in both BHC securities prices and BOPEC ratings. Detecting such a leading relationship between market variables and BOPEC ratings would support the notion that market signals could be useful to BHC supervisors. Failure to detect such a relationship could be interpreted as evidence that BHC assets are too opaque, thus rendering financial market signals of little use from a supervisory standpoint. Alternatively, failure to detect a leading relationship could imply that investors are indifferent to BOPEC rating changes.

Second, the timing convention used in the event study is critical. If investors receive information about a change in BHC condition before the beginning of our defined event window, then the event study will miss the market signal. To guard against this possibility, our event window begins one year before the BOPEC assignment. Additionally, we limit our analysis to events where there is at least one year between a BHC's inspections.

Third, care must be taken in interpreting signals extracted from market prices. For the case of subordinated debt spreads, market illiquidity and infrequent trading can make reported prices unreliable. These problems would tend to bias the results against finding a significant relationship between supervisory ratings and changes in spreads. For equity market data, our focus on stock returns abstracts from problems of disentangling changes in stock returns due to changes in the market value of BHC assets and those due to changes in BHC asset volatility.[12] These distinctions are important when trying to forecast events such as BHC default, especially in light of the public safety net for banks.

Finally, market signals can be interpreted as significant or abnormal only with the aid of a model. Thus, as with all event studies, the hypotheses tested below are actually joint hypotheses of whether market investors anticipate an event and whether the pricing model is correct. For our equity market event studies, we used a standard model. For our debt market event study, however, there is no commonly accepted model of bond spreads that allows a clear identification of BHC-specific variation. Rather, we look for significant deviations in BHC debt spreads from reasonable benchmarks.

The focus in this section on abnormal changes in security prices is based on the ease with which certain hypotheses can be tested in a univariate setting. This approach does not, of course, make use of the full potential of market data. If a common shock, such as a recession, affects all securities prices, our focus on abnormal returns would ignore this shock. Yet, the fact that all securities prices, including BHC security prices, have changed may be useful information to the supervisors. We address this issue of systematic variation in securities prices in our multivariate analysis in Section 5.

In both sets of event studies, we define the event date to be the inspection exit date, which is the day that examiners conduct their final meeting with BHC management. The first set of event studies is based on BHC stock returns and uses standard methods, as in Campbell, et al. (1997). In the second set, we examine whether changes in subordinated debt spreads anticipate changes in BOPEC ratings.

4.1. Equity Market Event Studies

For our equity market event study, we use monthly stock return data to investigate whether market investors anticipate BOPEC changes.[13] We define the event window as beginning twelve months prior to the exit date and ending three months after the exit date. We assume that stock returns follow a simple two-factor model, where the factors are the return on a broad market portfolio and the change in the federal funds rate; see Box 1 for further details. The model is estimated using monthly returns for up to 60 months prior to the beginning of the event window. In the event window, we compute *abnormal returns,* or deviations of actual returns from the model's predicted return. Cumulative abnormal returns (CARs) are formed by summing abnormal returns across the event window and are standardized using their estimated variances. The resulting standardized CARs (SCARs) are characterized well by the standard normal distribution and are used to test our event study hypotheses.

The event study is conducted using BOPEC ratings assigned from 1990 to 1999. A BHC inspection qualifies for the event study if it meets three criteria: there is enough stock return data to estimate the two-factor model; we can identify the BHC's lead bank and its prior BOPEC rating; and at least one year has elapsed since the previous inspection. The sample for this exercise consists of 813 BOPEC assignments for publicly traded BHCs, of which 139 (17 percent) are upgrades and 85 (10 percent) are downgrades.

In Figure 1, we plot the average CARs and SCARs for the BOPEC upgrades, downgrades, and no changes. The average CAR is meant to convey a sense of the economic impact of a BOPEC assignment on stock returns, while the average SCAR is meant to gauge whether this effect is sta-

12. We analyze this distinction explicitly in Krainer and Lopez (2001).

13. Note that we actually used four-week returns, but we use the term "monthly" for convenience.

Box 1
An Event Study Model of BHC Stock Returns

We structure our equity market event studies using a two-factor model of BHC stock returns; see Campbell, et al. (1997) for a general overview of event study methodology. We define R_{it} to be the four-week (or "monthly") stock return for BHC i at time t. R_{it} is assumed to take the form

$$R_{it} = \alpha + \beta R_{mt} + \gamma f_t + \varepsilon_{it},$$

where R_{mt} is the monthly return on the value-weighted portfolio of all NYSE, AMEX, and NASDAQ stocks at time t, f_t is the monthly change in the federal funds rate, and ε_{it} is a normally distributed error term.

In the banking literature, two-factor models are by far the most common specification used for event studies, and virtually all of these models incorporate the overall market return as a factor. There is less uniformity in the choice of the second factor. Some authors have used the return on a portfolio of bank stocks, as per Berger and Davies (1998). Other studies let the second factor capture changes in the interest rate environment. For example, Kwan (1991) used holding period returns on short-term and long-term constant maturity Treasury bonds as the second factor, and Hirtle (1997) used the percentage change in yield on a 10-year Treasury security. Following Kho, et al. (2000), we use the change in a short-term interest rate as the second factor.

We define the event window as beginning twelve months prior to the exit date of the BHC inspection and ending three months after the exit date. The model's parameters are estimated using monthly returns for up to 60 months prior to the beginning of the event window. Within the event window, the abnormal return is defined as

$$AR_{it} = R_{it} - \hat{\alpha}_i - \hat{\beta}_i R_{mt} - \hat{\gamma}_i f_t,$$

and the cumulative abnormal return from time $t-j$ to time t is defined as

$$CAR_{i,t-j,t} = \sum_{s=t-j}^{t} AR_{is}.$$

The standardized cumulative abnormal return for this period is calculated as

$$SCAR_{i,t-j,t} = \frac{CAR_{i,t-j,t}}{\sqrt{\text{var}(CAR_{i,t-j,t})}},$$

which has a t-distribution with degrees of freedom based on the size of the model's estimation window. These *SCAR* variables are the ones used to test our event study hypotheses.

Note that we use the fitted systematic returns from our two-factor model in our multivariate analysis. The fitted systematic return is defined as

$$SR_{it} = \hat{\alpha}_i + \hat{\beta}_i R_{mt} + \hat{\gamma}_i f_t,$$

and the cumulative systematic return from time $t-j$ to time t is defined as

$$CSR_{i,t-j,t} = \sum_{s=t-j}^{t} SR_{is}.$$

The standardized cumulative systematic return for this period is calculated as

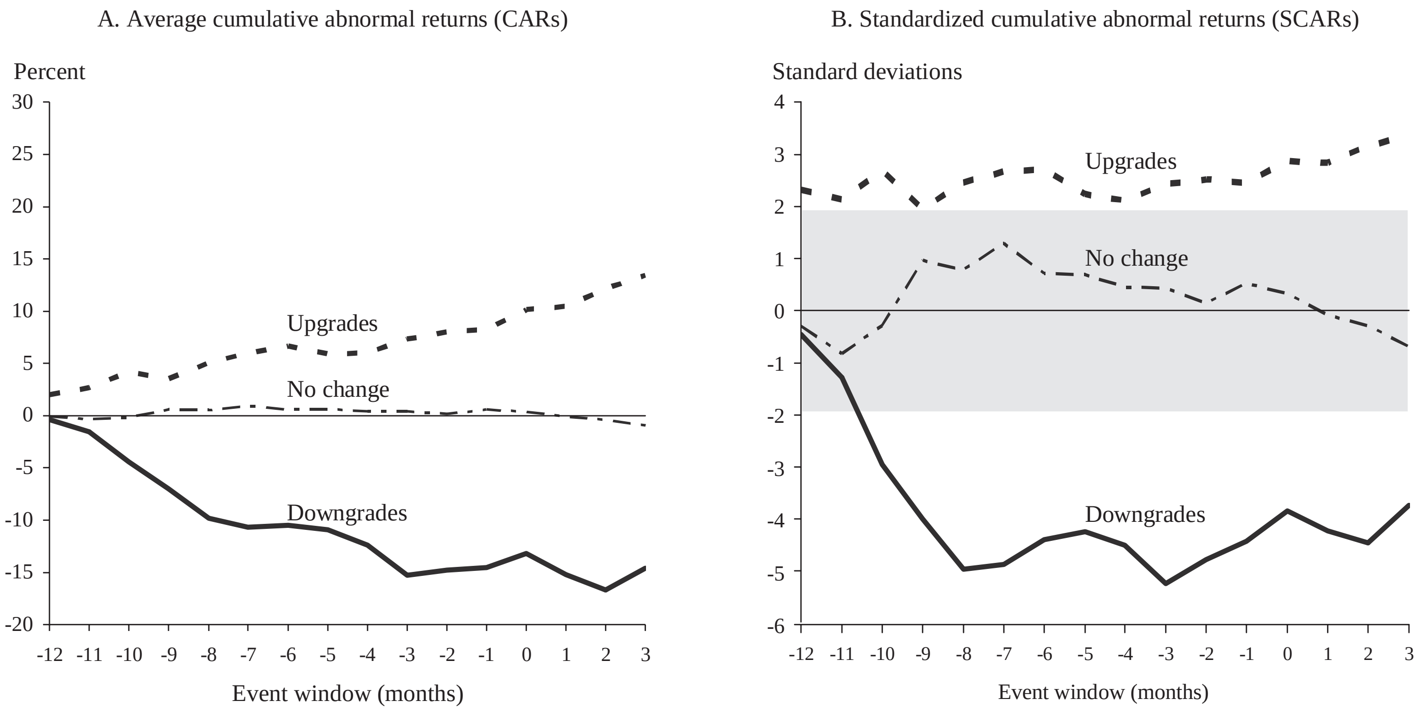
$$SCSR_{i,t-j,t} = \frac{CSR_{i,t-j,t}}{\sqrt{\text{var}(CAR_{i,t-j,t})}}.$$

tistically significant. Note that average SCARs with an absolute value greater than 1.96 are statistically significant at the 5 percent confidence level.

The figure clearly shows that equity market participants anticipate BOPEC changes well in advance of their actual assignments. For the BOPEC upgrades, the average CARs increase as we approach the event date, growing to over 13 percent by the end of the event window. The SCARs are significantly positive as early as twelve months prior to the inspection. For the BOPEC downgrades, the average CAR was –15 percent by the end of the event window, and the average downgrade SCARs are statistically significant as early as ten months prior to the exit date. Additionally, on average, the market does not appear to send false signals of approaching changes. The average CARs for inspections with no change in BOPEC rating are generally positive yet quite small and never statistically significant. By the end of the event window, the average CAR for no change inspections is approximately –1 percent. These results confirm the hypothesis that equity markets are capable of anticipating BOPEC changes.

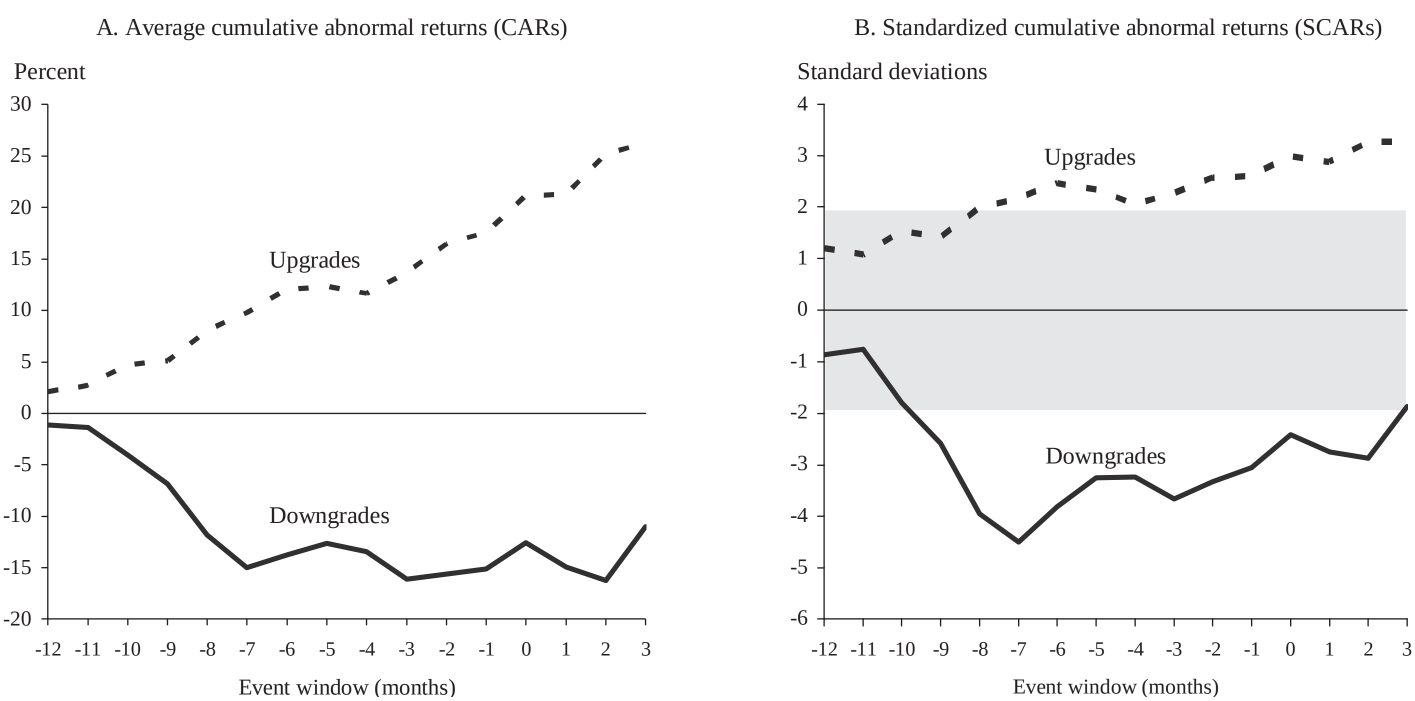
A change over the threshold between BOPEC ratings 2 and 3 is of particular concern to supervisors, and our event study focusing on these occasions is summarized in Figure 2. Note that crossings of this regulatory threshold are relatively rare events (about 11 percent of the sample) and, thus, the average CARs and SCARs are measured with more error. Nevertheless, the stock market behavior prior to these events is broadly similar to our earlier results. For the 39 downgrades below the threshold, the market sends a statistically significant signal nine months prior to the exit date, approximately the same time as for the case of all downgrades. Over the event window, the average CAR for downgrades below the threshold is –11 percent.

FIGURE 1
EQUITY MARKET EVENT STUDY FOR BOPEC CHANGES



Note: Inspection exit date occurs at month 0. Sample consists of 813 inspections, of which 139 are upgrades, 85 are downgrades, and 589 show no change. Shaded band shows the 95 percent confidence interval.

FIGURE 2
EQUITY MARKET EVENT STUDY FOR BOPEC CHANGES CROSSING THE THRESHOLD BETWEEN RATINGS 2 AND 3



Note: Inspection exit date occurs at month 0. Sample consists of 50 upgrades and 39 downgrades over the BOPEC 2 and BOPEC 3 threshold. Shaded band shows the 95 percent confidence interval.

For the sample of 50 upgrades above the threshold, the average CAR is 21 percent at the event date and 26 percent by the end of the event window. The equity market appears to anticipate BOPEC upgrades moving above the threshold by seven months. This more focused event study confirms that equity markets are capable of anticipating BOPEC changes and that equity market variables could be useful for supervisory BHC monitoring.

4.2. Debt Market Event Study

In this section, we investigate whether changes in BHC bond yields anticipate changes in supervisory BOPEC ratings. The debt market variables used in this study are taken from the Warga/Lehmann Brothers Corporate Bond Database and are the same as those used by Bliss and Flannery (2001).[14]

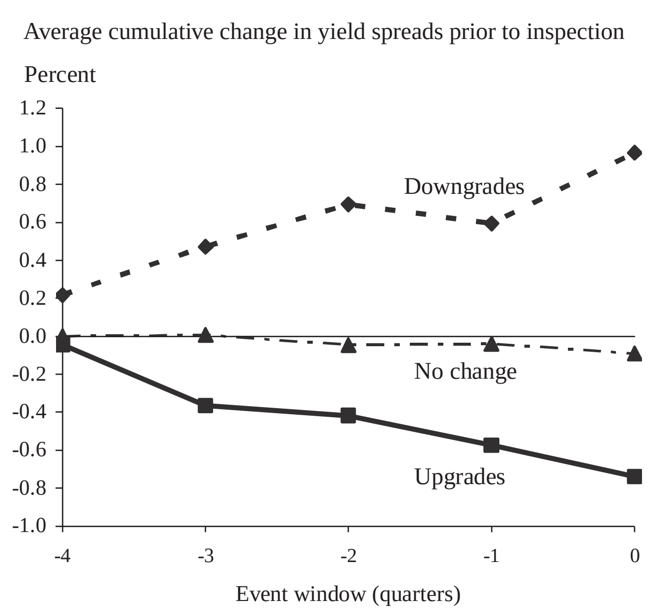
There are two unique empirical issues regarding bond data that must be addressed. First, in cases where a BHC has multiple outstanding bonds, we must compress these multiple market signals into a single observation. When confronted with this problem in our data set, we use a weighted average change in bond yields, where the weights are the bond amounts outstanding in each quarter. Second, as with the equity market variables, it is necessary to have some measure of whether a change in bond yield is abnormal. Following Bliss and Flannery (2001), we create BHC yield spreads by subtracting from a BHC's yield the yield on an index of bonds with similar terms-to-maturity and agency ratings. We use indices based on eleven categories corresponding to Moody's and S&P ratings and on three term-to-maturity categories.[15] This adjustment, however, still leaves open the possibility that the yield spread contains information common to all BHCs. Therefore, we further adjust the yield spread by subtracting off the median yield spread of BHC bonds with similar BOPEC ratings. The resulting variable is our "adjusted yield spread."

Due to data availability, the sample used in this event study is smaller than the one used in the equity event studies. This loss of observations reflects the relative scarcity of BHCs with publicly traded debt compared to those with publicly traded equity. This subsample contains 315 BOPEC assignments for 63 unique BHCs, 57 (18 percent) of which are upgrades and 30 (10 percent) of which are downgrades. The sample period ranges from the first quarter of 1990 to the second quarter of 1998.

In anticipation of a BOPEC downgrade, we would expect an increase in the BHC's adjusted yield spread, and in anticipation of an upgrade, we would expect a decrease in the adjusted yield spread. As shown in Figure 3, the data appear to move in accordance with these expectations. For the BOPEC downgrades, the average cumulative increase in adjusted yield spreads is about 100 basis points (i.e., 1%) by the time of the exit date. For the BOPEC upgrades, adjusted yield spreads drop by about 75 basis points by the exit date. Using the nonparametric sign test, the cumulative change in adjusted yield spreads for the BOPEC upgrades and downgrades are significantly different from zero at the 5 percent confidence level. For the no-change inspections, the average cumulative change in adjusted yield spread is about –5 basis points and is not statistically significant.

In conclusion, both sets of event studies suggest that, on average, changes in BHC stock returns and subordinated debt yields are consistent with the supervisory BHC assessments summarized by BOPEC ratings. These financial market signals can be observed as statistically significant from about twelve to nine months prior to the BOPEC assignments, and thus these data could be of use for supervisory off-site monitoring. What remains to be seen, however, is whether the financial variables add information to that which is already available to supervisors from their own information sources. We turn to this question in the next section.

Figure 3
Debt Market Event Study for BOPEC Changes



Note: Sample consists of 315 inspections, of which 57 are upgrades, 30 are downgrades, and 228 show no change.

14. We thank Rob Bliss for sharing these data with us.

15. As in Bliss and Flannery (2001), the qualifiers attached to the basic agency ratings are suppressed, and the three maturity categories are less than five years, five to ten years, and greater than ten years.

5. Multivariate Analysis Using the BOPEC Off-Site Monitoring Model

The second policy question that we address is whether financial market data provide supervisors with information above and beyond the balance sheet and supervisory ratings data that they already possess. Unlike the event studies where we looked at the behavior of securities prices leading up to BOPEC assignments of known outcomes, here we formally model the supervisory BOPEC ratings themselves. While models must abstract from the real decisionmaking process that takes place in an inspection, they have the advantage of offering a relatively easy way to evaluate the potential contribution of different data.

5.1. Model Description

Following Krainer and Lopez (2001, 2003), we use a multivariate model that allows BOPEC ratings to depend on both supervisory variables and financial market data. Not only does this approach allow us to study the marginal value of financial market data relative to supervisory data, but we are also free to use a larger array of financial market data than is possible in a univariate framework; specifically, we can incorporate both equity and debt market data in the same model. We also now can examine the influence of systematic changes in financial market variables (i.e., variation across all BHCs and not limited to a specific BHC) on BOPEC ratings.

The BOPEC off-site monitoring (BOM) model that we use is an ordered logit; see Box 2 for further details. In the model, the BOPEC rating assigned to a BHC depends on supervisory data available two quarters prior to the inspection and on financial market variables available one quarter prior to the inspection. We would prefer to use more up-to-date supervisory variables for this exercise, but, as a practical matter, supervisory data are not widely available for about 60 days after the end of the quarter. Since we conduct a forecasting exercise, we err on the conservative side and use supervisory data from two quarters prior to the inspection so as to best mimic the information supervisors actually would have if they were generating the forecasts in real time.[16]

Supervisors collect hundreds of financial variables in the course of routine BHC monitoring and, for the case of the largest BHCs where examiners are continuously present, they collect data more often than the quarterly frequency that we use here. To build a parsimonious off-site monitoring model, we condense the list of potential explanatory variables to just eight variables that correspond to the main areas of concern in BHC inspections.

To capture the supervisory concerns regarding bank subsidiaries summarized in the "B" component of the rating, we use the CAMELS rating of the BHC's lead bank; the ratio of the sum of BHC nonperforming loans, nonaccrual loans, and other real estate owned to its total assets; and the ratio of BHC allowances (or provisions) for losses on loans and leases to its total loans. To measure a BHC's nonbank activities that are captured in the "O" component of the rat-

Box 2
THE BOPEC OFF-SITE MONITORING (BOM) MODEL

As proposed by Krainer and Lopez (2001), an ordered logit model is used to construct the BOPEC off-site monitoring (BOM) model. Within this structure, BOPEC ratings are modeled as a continuous variable BP_{it}^*, which is a linear function of supervisory and financial market variables. Specifically, $BP_{it}^* = \beta x_{it-2} + \theta z_{t-1} + \varepsilon_{it}$, where x_{it-2} is a vector of supervisory variables unique to BHC i observed two quarters prior to the inspection, $z_{i,t-1}$ is a vector of financial market variables specific to BHC i observed one quarter prior to the inspection, and the disturbance term ε_{it} has a standard logistic distribution. Note that this is a forecasting model, so all the independent variables are lagged relative to BP_{it}^*. We would prefer to use more up-to-date supervisory variables for this exercise, but, as a practical matter, supervisory data are not widely available for about 60 days after the end of the quarter. Thus, we must use supervisory data from two quarters prior to the inspection.

Since BOPEC ratings only take integer values between 1 and 5, the model's estimation procedure also must find the four points, denoted α_1 through α_4, that cut the support of BP_{it}^* into five non-overlapping regions. That is,

$$BP_{it} = 1 \text{ if } BP_{it}^* \in (-\infty, \alpha_1],$$

$$BP_{it} = 2 \text{ if } BP_{it}^* \in (\alpha_1, \alpha_2],$$

$$BP_{it} = 3 \text{ if } BP_{it}^* \in (\alpha_2, \alpha_3],$$

$$BP_{it} = 4 \text{ if } BP_{it}^* \in (\alpha_3, \alpha_4],$$

$$BP_{it} = 5 \text{ if } BP_{it}^* \in (\alpha_4, \infty).$$

The model is estimated using maximum likelihood methods. The estimation results are available in Krainer and Lopez (2001, 2003).

16. As a robustness check, we verify that the results change little when the supervisory data are lagged by only one quarter.

ing, we use an indicator of whether the BHC engages in securities underwriting via a Section 20 subsidiary and the ratio of its trading assets to total assets. These nonbank activities could affect BHC condition in a number of different and competing ways. For example, if a BHC's expansion into nonbanking activities affords it revenue diversification, the expansion could improve the BOPEC rating, holding all other things equal. Alternatively, such activities could increase overall BHC risk, in which case the BOPEC rating would be harmed.

The health of the parent company, as reflected in the "P" component of the rating, is captured by the ratio of the lead bank's equity capital to the BHC's equity capital—the so-called "double leverage" ratio. The earnings (or "E") component of the rating is captured using the BHC's return on average assets, defined as the ratio of the four-quarter average of a BHC's net income to the four-quarter average of its assets. The BHC's capital position, which is the "C" component of the rating, is summarized by the ratio of BHC equity capital to its total assets.

We also include two BHC control variables in the regression. We include the natural log of total BHC assets as a size control. We also include the BHC's last assigned BOPEC rating as a way to summarize any relevant supervisory information not captured by the other eight variables.

For financial market variables, we use the equity and debt variables from the event studies. For the equity variables, we use the SCAR variables over the six-month window at one quarter prior to the BOPEC assignment. We also include a measure of the overall equity market variation within the BHC stock returns. This latter variable is used within the model to capture broad economic shocks that likely will affect all BHC securities prices.[17] The debt market variable is calculated as the three-month change in the adjusted yield spread ending one quarter prior to the exit date.

The full estimation sample contains 2,940 inspections beginning in the first quarter of 1990 and ending in the second quarter of 1998, of which 643 (22 percent) are upgrades and 344 (12 percent) are downgrades. We consider two specifications of the BOM model. The first model has only supervisory variables and is known as the "core" model. The second model is an extension of the core model that includes the three financial market variables. When estimating the extended model, the data are pooled to include all BHCs, regardless of whether they have publicly traded equity or publicly traded debt.[18]

5.2. *In-Sample Estimation Results*

In-sample estimates of the BOM model's coefficients over the full estimation sample are discussed in great detail in Krainer and Lopez (2001, 2003) and are summarized briefly in Box 3. Key supervisory variables, such as the lagged BOPEC rating, the bank subsidiary's CAMELS rating, problem loans, provisions, returns on assets, and equity capital, all have coefficients with the expected signs and are statistically significant. Thus, using the BOM model, the variables meant to proxy for the "B," "E," and "C" components of the BOPEC ratings are significant for describing actual ratings. However, our measures of the "O" and "P" components are not.

The financial market variables all have the expected sign and are statistically significant. Thus, both equity and debt market information appear to be useful for describing past BOPEC ratings, suggesting that both sets of market participants have information that could be useful to supervisors. For the equity market variables, both the BHC-specific return and the systematic return are significantly different from zero. Thus, as hypothesized, systematic changes in stock prices that would accompany common shocks, such as changing economic conditions, contain information that could be useful to supervisors. Large abnormal changes in stock prices and debt yields again are shown to contain useful information, as was found in the event studies.

5.3. *Out-of-Sample Forecast Accuracy*

Although the in-sample estimation results are of interest, what is of most interest for an off-site monitoring model is its out-of-sample forecast accuracy. That is, given a model fitted to historical data, what percentage of its forecasted BOPEC ratings correctly anticipate the BOPEC ratings eventually assigned? To address this question, we generate BOPEC forecasts from the core and extended BOM models and assess their performance relative to the actual BOPEC assignments in our database. We do so by using a rolling estimation technique; that is, we estimate the two BOM models using just four quarters of data and then forecast BOPEC ratings for the next quarter. We roll through our sample period from the first quarter of 1990 through

17. By including these two equity market variables in the model, we effectively decompose the cumulative BHC stock returns into a fitted systematic component and an abnormal (or idiosyncratic) component; see Box 1 as well as Krainer and Lopez (2001) for further details.

18. We pool the data and adjust for the missing securities market information in order to increase our sample size and improve the precision of the coefficient estimates. See Krainer and Lopez (2001, 2003) for further methodological details.

Box 3
Explanatory Variables in the Extended BOM Model

Variables	An increase suggests
Supervisory variables:	
Lead bank CAMELS	worse rating*
Problem loans / total assets	worse rating*
Provisions / total loans	worse rating*
Section 20 indicator	unclear
Trading assets / total assets	unclear
Double leverage	improved rating
Return on assets	improved rating*
Equity capital	improved rating*
Log of total assets	unclear*
Prior BOPEC rating	worse rating*
Financial market variables:	
Six-month SCAR	improved rating*
Six-month fitted return	improved rating*
Three-month change in adjusted yield spread	worse rating*

*Indicates statistical significance at the 5 percent level.

the second quarter of 1998 and generate one-quarter-ahead forecasts for all 2,940 inspections in the database for each quarter. To analyze forecast accuracy, we look at the BOPEC forecasts generated four quarters prior to the BOPEC assignments.

The results are summarized in Tables 2 and 3, where a row describes the distribution of actual BOPEC ratings associated with a particular BOPEC forecast. As we can see along the tables' diagonals, both models do fairly well at forecasting future ratings even at four quarters prior to the inspection. For example, in Table 2, if the core model forecasts a BOPEC rating of 1 four quarters prior to an inspection, the actual BOPEC rating assigned is 1 about 80 percent of the time. The core model forecasts a BOPEC rating of 1 or 2 accurately 75 percent of the time. However, the model appears to be less accurate in forecasting inspection outcomes at the lower-rated BHCs at the four-quarter horizon. Given BOPEC forecasts of 3 or higher, these forecasts are correct about 40 percent of the time.

The extended BOM model generates forecasts with approximately the same degree of forecast accuracy. As shown in Table 3, if the extended BOM model generates a BOPEC forecast of 1, these forecasts again are correct 78 percent of the time. BOPEC forecasts of 1 or 2 again are correct about 75 percent of the time. Like the core model, the extended model's accuracy diminishes for lower rated BHCs. BOPEC forecasts of 3 are correct only 41 percent of the time and actually are BOPEC ratings of 2 approximately 47 percent of the time. BOPEC forecasts of 5 match the actual rating just 19 percent of the time and are actually 4 ratings 53 percent of the time.

This poor performance by both models at the lower end of the ratings distribution may be due to the fact that the bulk of the BOPEC ratings in our sample are 1 and 2 (recall

Table 2
Forecasting Accuracy of the Core Model at Four Quarters Prior to Assignment

	Actual rating					
Forecasted rating	1	2	3	4	5	Total
1	**619**	148	4	1	1	773
% of total	**80**	19	1	0	0	
2	302	**1,138**	116	26	1	1,583
% of total	19	**72**	7	2	0	
3	1	180	**153**	39	0	373
% of total	0	48	**41**	11	0	
4	0	16	70	**76**	11	173
% of total	0	9	41	**44**	6	
5	0	0	9	20	**9**	38
% of total	0	0	24	53	**24**	
Total	922	1,482	352	162	22	2,940

Note: Rows denote the distribution of actual BOPEC ratings given the model's forecast. The full estimation sample contains 2,940 inspections beginning in the first quarter of 1990 and ending in the second quarter of 1998, of which 643 (22 percent) are upgrades and 344 (12 percent) are downgrades. Bold denotes correct forecast.

Table 3
Forecasting Accuracy of the Extended Model at Four Quarters Prior to Assignment

	Actual rating					
Forecasted rating	1	2	3	4	5	Total
1	**605**	161	5	1	1	773
% of total	**78**	21	1	0	0	
2	316	**1,131**	116	22	0	1,585
% of total	20	**71**	7	1	0	
3	1	171	**152**	42	1	367
% of total	0	47	**41**	11	0	
4	0	19	69	**78**	13	179
% of total	0	11	39	**44**	7	
5	0	0	10	19	**7**	36
% of total	0	0	28	53	**19**	
Total	922	1,482	352	162	22	2,940

Note: Extended model includes supervisory data and debt market and equity market data as explanatory variables. Rows denote the distribution of actual BOPEC ratings given the model's forecast. The full estimation sample contains 2,940 inspections beginning in the first quarter of 1990 and ending in the second quarter of 1998, of which 643 (22 percent) are upgrades and 344 (12 percent) are downgrades. Bold denotes correct forecast.

Table 1). Another possible reason is that BOPEC ratings of 3 and 5 are important cusp points in the rating system and may be where supervisory concerns not captured by the models play a larger role.

A critical question is whether the extended model including both supervisory and financial market variables provides useful information about BOPEC ratings beyond what is obtained by the core model. A common way to make such an assessment is to compare the accuracy of the two sets of forecasts statistically, which in this case is comparing the percentages of BOPEC ratings that they accurately forecast. We formally test for differences between the forecasts in other research (Krainer and Lopez 2001, 2003) and we find little statistical difference between the two sets of forecasts. Hence, using this metric, the financial market variables do not appear to contribute to the supervisory information set.

This result, however, does not mean that the set of BOPEC ratings correctly forecasted by the two models are identical; in fact, the two models signal BOPEC changes for different, although overlapping, sets of BHCs. The forecasting literature has shown that combining forecasts from different models can improve certain aspects of forecast accuracy. Hence, another way to gauge the contribution of financial market information to the BOM model is to examine the additional forecast signals for publicly traded BHCs generated by the extended model relative to the core model's signals. Seen in this light, the marginal benefit of adding these signals to those from the core model is notable.

In the second column of Table 4, we focus exclusively on downgrade signals prior to assignment. We define a downgrade signal as a forecasted BOPEC rating that is greater than the current rating by one or more ratings.[19] Using the signals generated by both models, we ask, what is the percentage increase in correct signals when financial market data are used in the BOM model? For the extended model, 9 percent more correct signals are produced at the four-quarter horizon over and above those produced by the core model.[20] At the one-quarter horizon, the model produces 37 percent more signals.

Of course, the extended model produces incorrect signals over and above those produced by the core model. We look at this tradeoff more closely in the third column of Table 4, where we express the ratio of correct signals to incorrect signals. For example, the extended model produces one additional correct downgrade signal at the cost of four incorrect signals at the four-quarter horizon. At the one-quarter horizon, however, the accuracy tradeoff dramatically improves to four additional correct signals at the cost of only one extra incorrect signal. In summary, the marginal contribution of financial market variables to BOPEC forecasts generated by the extended BOM model is present, but not across all BOPEC forecasts. As presented in Table 4, the contribution is most clearly seen in the additional correct forecasts generated at the relatively low cost of incorrect forecasts.

This result leads us directly to our third public policy question about the degree of accuracy supervisors should demand of financial market signals and of off-site monitoring models more generally. Clearly, any model will generate correct and incorrect forecasts, and the aggregate value of these forecasts will depend on the supervisors' relative costs regarding these forecasts. The two types of errors that can be made by an off-site monitoring model are missed signals, which are occasions when observed BOPEC downgrades are not forecasted (i.e., type-1 error), and false positives, which are forecasted downgrades that do not occur (i.e., type-2 error). Supervisors should be aware of the tradeoffs between these two error types for a given set of forecasts and could incorporate the tradeoff costs directly into their forecast evaluation.

For our exercise in Table 4, at four quarters prior to assignment, we have one missed downgrade signal for every four false positives. The policy question regarding the use of financial market information is whether the benefit of catching an additional BOPEC downgrade one year before

Table 4
Forecast Analysis of the Combined Forecasts from the Core and Extended BOM Models

	Extended BOM Model	
	% increase in total correct downgrade forecast signals	Tradeoff of correct signals to incorrect signals
4 quarters	9	1 / 4
3 quarters	38	5 / 2
2 quarters	51	19 / 3
1 quarter	37	4 / 1

Notes: Downgrade signal is defined as (forecasted rating – current rating) > 1. Table reports the number of downgrades correctly signaled by the extended model and not identified by the core model, expressed as a percentage of downgrades correctly identified by the core model. Entries in the tradeoff column report the number of additional correct BOPEC downgrade signals from the extended BOM model over the core model, at the cost of additional incorrect downgrade signals.

19. We set our definition of a downgrade signal at a forecast difference of one rating for clarity. Weaker signals, such as a difference of 0.75, could be used with a concurrent increase in the total number of signals and in the number of incorrect signals.

20. Formally, we compute the number of correct extended BOM signals and subtract those downgrades that the core BOM also correctly identifies.

it happens is worth the cost of four additional on-site inspections that do not lead to downgrades. This tradeoff becomes more stark at one quarter prior to assignment. Without knowing these actual costs a priori, it is not possible to make this judgment for our example. However, given the potentially large costs of missed signals, it is clear that supervisors could find it useful to use the extended BOM model for off-site monitoring.

6. Conclusion

Banks and their holding companies are subject to government supervision and regulation to guard against systemic risk to the economy. These supervisory efforts use a wide variety of information sources, from on-site inspections to BHC securities prices. In this article, we address three public policy questions related to the use of financial market information for supervisory purposes and conclude that this information should be of use to supervisors, at least in their off-site monitoring efforts.

The first policy question is whether changes in BHC risk characteristics get incorporated into BHC securities prices, the primary source of financial market information. In agreement with much of the academic literature, we find that securities prices do incorporate these risk changes and that the financial markets anticipate supervisory rating changes by about nine to twelve months.

The second question is whether supervisors already capture the information content of financial market prices using other means. Using the BOPEC off-site monitoring (BOM) model proposed by Krainer and Lopez (2001), we find that both equity and debt market variables are statistically significant in explaining BOPEC rating assignments, even when a wide variety of supervisory variables are already in the model. This result is further evidence that supervisors could benefit from incorporating financial market information into their off-site monitoring efforts.

The final policy question is how best to evaluate the contribution of financial market variables to supervisory monitoring; that is, what degree of accuracy should supervisors demand of financial market signals and of off-site monitoring models more generally? To address this question, we focus on out-of-sample forecast accuracy. We find that the forecast accuracy of BOM models with and without financial market variables is not statistically different, but that the two sets of BHCs for which BOPEC changes are forecasted are not identical. Furthermore, we find that when the downgrade forecast signals from the two models are combined, we have a roughly 10 percent increase in the number of correct signals, although at the cost of increased false positive signals. The policy question regarding the use of financial market information is now whether the benefit of catching an additional BOPEC downgrade one year before it happens is worth the cost of four additional on-site inspections that do not lead to downgrades. Given the potentially large costs of missed signals of BOPEC changes, especially downgrades, we believe it is clear that supervisors could find it beneficial to use the extended BOM model and financial market information more broadly for off-site monitoring.

References

Allen, L., J. Jagtiani, and M. Moser. 2001. “Further Evidence on the Information Content of Bank Examination Ratings: A Study of BHC-to-FHC Conversion Applications.” *Journal of Financial Services Research* 20, pp. 213–232.

Avery, R.B., T.M. Belton, and M.A. Goldberg. 1988. “Market Discipline in Regulating Bank Risk: New Evidence from the Capital Markets.” *Journal of Money, Credit, and Banking* 20, pp. 597–610.

Basel Committee on Banking Supervision. 2001. “The New Basel Accord.” Consultative Paper, Bank for International Settlements. http://www.bis.org/publ/bcbsca.htm

Berger, A.N., and S.M. Davies. 1998. “The Information Content of Bank Examinations.” *Journal of Financial Services Research* 14, pp. 117–144.

Berger, A.N., S.M. Davies, and M.J. Flannery. 2000. “Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When?” *Journal of Money, Credit, and Banking* 32, pp. 641–667.

Bliss, R. 2000. “The Pitfalls in Inferring Risk from Financial Market Data.” Working Paper 2000-24, FRB Chicago.

Bliss, R. and M. Flannery. 2001. “Market Discipline in the Governance of U.S. Bank Holding Companies: Monitoring versus Influence.” In *Prudential Supervision: Why Is It Important and What Are the Issues?* ed. R. Mishkin. Cambridge, MA: NBER.

Board of Governors of the Federal Reserve System. 1995. “Revised Bank Holding Company Surveillance Procedures.” Supervisory Letter 95-43. http://www.federalreserve.gov/boarddocs/SRLETTERS/1995/sr9543.htm

Board of Governors. 1999. “Using Subordinated Debt as an Instrument of Market Discipline: Report of the Study Group on Subordinated Notes and Debentures.” Federal Reserve System Staff Study 172. http://www.federalreserve.gov/pubs/staffstudies/1990-99/172sum.htm

Board of Governors. 2001. “A User’s Guide for the Bank Holding Company Performance Report.” http://www.federalreserve.gov/boarddocs/supmanual/default.htm#bhcuser

Board of Governors. 2002. “Revisions to Bank Holding Company Supervision Procedures for Organizations with Total Consolidated Assets of $5 Billion or Less,” Supervisory Letter 02-01. http://www.federalreserve.gov/boarddocs/SRLETTERS/2002/sr0201.htm

Board of Governors and the U.S. Department of the Treasury. 2000. “The Feasibility and Desirability of Mandatory Subordinated Debt,” Report to Congress pursuant to Section 108 of the Gramm-Leach-Bliley Act of 1999.

Bongini, P., L. Laeven, and G. Majnoni. 2002. “How Good Is the Market at Assessing Bank Fragility? A Horse Race between Different Indicators.” *Journal of Banking and Finance* 26, pp. 1,011–1,028.

Campbell, J., A. Lo, and A. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ : University Press.

Cole, R.A., B.G. Cornyn, and J.W. Gunther. 1995. “FIMS: A New Monitoring System for Banking Institutions.” *Federal Reserve Bulletin* (January) pp. 1–15.

Cole, R.A., and J.W. Gunther. 1998. “Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems.” *Journal of Financial Services Research* 13, pp. 103–117.

Covitz, D.M., D. Hancock, and M.L. Kwast. 2002. “Market Discipline in Banking Reconsidered: The Role of Deposit Insurance Reform, Funding Manager Decisions, and Bond Market Liquidity.” Manuscript, Division of Research and Statistics, Board of Governors of the Federal Reserve System.

Curry, T.J., P.J. Elmer, and G. Fissel. 2001. “Regulator Use of Market-Related Data to Improve the Identification of Bank Financial Health.” Manuscript, Federal Deposit Insurance Corporation.

Dahiya, S., A. Saunders, and A. Srinivasan. 2003. “Financial Distress and Bank Lending Relationships.” *Journal of Finance* 58, pp. 375–399.

DeFerrari, L., and D. Palmer. 2001. “Supervision of Large Complex Banking Organizations.” *Federal Reserve Bulletin* (February) pp. 47–57.

DeYoung, R., M.J. Flannery, W.W. Lang, and S.M. Sorescu. 2001. “The Information Content of Bank Exam Ratings and Subordinated Debt Prices.” *Journal of Money, Credit, and Banking* 33(4) (November) pp. 900–925.

Diamond, D.W. 1984. “Financial Intermediation and Delegated Monitoring,” *Review of Economic Studies* 51, pp. 393–414.

Elmer, P.J., and G. Fissel. 2001. “Forecasting Bank Failure from Momentum Patterns in Stock Returns.” Manuscript, Federal Deposit Insurance Corporation.

Estrella, A., S. Park, and S. Peristiani. 2000. “Capital Ratios as Predictors of Bank Failure.” FRB New York *Economic Policy Review* (July) pp. 33–52.

Evanoff, D., and L. Wall. 2001. “Sub-Debt Yield Spreads as Bank Risk Measures.” *Journal of Financial Services Research* 20, pp. 121–145.

Flannery, M.J. 1998. “Using Market Information in Prudential Bank Supervision.” *Journal of Money, Credit, and Banking* 30(3) (August) pp. 273–305.

Flannery, M.J. 2001. “The Faces of ‘Market Discipline.’” *Journal of Financial Services Research* 20, pp. 107–119.

Flannery, M., S. Kwan, and M. Nimalendran. Forthcoming. “Market Evidence on the Opaqueness of Banking Firms’ Assets.” *Journal of Financial Economics*.

Flannery, M., and S. Sorescu. 1996. “Evidence of Bank Market Discipline in Subordinated Debenture Yields: 1983–1991.” *Journal of Finance* 51, pp. 1,347–1,377.

Gilbert, A., A. Meyer, and M. Vaughn. 2001. “Can Feedback from the Jumbo CD Market Improve Off-Site Surveillance of Small Banks?” Manuscript, Economic Research Department, FRB St. Louis.

Gorton, G., and A.M. Santomero. 1990. “Market Discipline and Bank Subordinated Debt.” *Journal of Money, Credit, and Banking* 22, pp. 119–128.

Gropp, R., and A.J. Richards. 2001. “Rating Agency Actions and the Pricing of Debt and Equity of European Banks: What Can We Infer about Private Sector Monitoring of Bank Soundness?” Working Paper #76, European Central Bank.

Gropp, R., J. Vesala, and G. Vulpes. 2002. "Equity and Bond Market Signals as Leading Indicators of Bank Fragility," Working Paper 150, European Central Bank.

Gunther, J.W., M.E. Levonian, and R.R. Moore. 2001. "Can the Stock Market Tell Bank Supervisors Anything They Don't Already Know?" FRB Dallas *Economic and Financial Review,* Second Quarter, pp. 2–9.

Gunther, J.W., and R.R. Moore. 2000. "Early Warning Models in Real Time." Financial Industry Studies Working Paper 1-00, FRB Dallas.

Hall, J.R., T.B. King, A.P. Meyer, and M.D. Vaughan. 2001. "What Can Bank Supervisors Learn from the Equity Markets? A Comparison of the Factors Affecting Market-Based Risk Measures and BOPEC Scores." Manuscript, FRB St. Louis.

Hancock, D., and M. Kwast. 2001. "Using Subordinated Debt to Monitor Bank Holding Companies: Is It Feasible?" *Journal of Financial Services Research* 20, pp. 147–187.

Hannan, T., and G.A. Hanweck. 1988. "Bank Insolvency Risk and the Market for Large Certificates of Deposit." *Journal of Money, Credit, and Banking* 20, pp. 203–212.

Hirtle, B. 1997. "Derivatives, Portfolio Composition, and Bank Holding Company Interest Rate Risk Exposure." *Journal of Financial Services Research* 12, pp. 243–266.

Hirtle, B., and J.A. Lopez. 1999. "Supervisory Information and the Frequency of Bank Examinations." FRB New York *Economic Policy Review* (April) pp. 1–20.

Jagtiani, J., G. Kaufman, and C. Lemieux. 2000. "Do Markets Discipline Banks and Bank Holding Companies? Evidence from Debt Pricing." Emerging Issues Series #99-3R, Supervision and Regulation, FRB Chicago.

Kho, B.C., D. Lee, and R.M. Stulz. 2000. "U.S. Banks, Crises, and Bailouts: From Mexico to LTCM." *American Economic Review Papers and Proceedings* 90, pp. 28–31.

Korobow, L., D.P. Stuhr, and D. Martin. 1977. "A Nationwide Test of Early Warning Research in Banking." FRB New York *Quarterly Review* (August) pp. 37–52.

Krainer, J., and J.A. Lopez. 2001. "Incorporating Equity Market Information into Supervisory Monitoring Models." FRB San Francisco Working Paper 2001-14.

Krainer, J., and J.A. Lopez. 2003. "Forecasting Supervisory Ratings Using Securities Market Information." Manuscript, Economic Research Department, FRB San Francisco.

Kwan, S.H. 1991. "Re-examination of Interest Rate Sensitivity of Commercial Bank Stock Returns Using a Random Coefficient Model." *Journal of Financial Services Research* 5, pp. 61–76.

Morgan, D.P. 2003. "Rating Banks: Risk and Uncertainty in an Opaque Industry." *American Economic Review* 92, pp. 874–888.

Morgan, D.P., and K.J. Stiroh. 2001. "Market Discipline of Banks: The Asset Test." *Journal of Financial Services Research* 20, pp. 195–208.

Pettway, R.H. 1976. "Market Tests of Capital Adequacy of Large Commercial Banks." *Journal of Finance* 31, pp. 865–875.

Sahajwala, R., and P. Van der Bergh. 2000. "Supervisory Risk Assessment and Early Warning Systems." Basel Committee on Banking Supervision Working Paper No. 4, Bank for International Settlements.

Stuhr, D.P., and R. Van Wicklen. 1974. "Rating the Financial Condition of Banks: A Statistical Approach to Aid Bank Supervisors." FRB New York *Monthly Review* (September) pp. 233–238.